



Nikolov, SG., Bull, DR., Canagarajah, CN., Halliwell, M., & Wells, PNT. (2000). Fusion of 2-D images using their multiscale edges. In *15th International Conference on Pattern Recognition (ICPR), Barcelona, Catalonia, Spain* (Vol. 3, pp. 41 - 44). Institute of Electrical and Electronics Engineers (IEEE).  
<https://doi.org/10.1109/ICPR.2000.903481>

Peer reviewed version

Link to published version (if available):  
[10.1109/ICPR.2000.903481](https://doi.org/10.1109/ICPR.2000.903481)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Fusion of 2-D Images Using Their Multiscale Edges

S. G. Nikolov, D. R. Bull, C. N. Canagarajah

Image Communications Group

Centre for Communications Research

University of Bristol, Merchant Venturers Building

Woodland Road, Bristol BS8 1UB, UK

{Stavri.Nikolov,Dave.Bull,Nishan.Canagarajah}@bristol.ac.uk

M. Halliwell, P. N. T. Wells

Centre for Physics and Engineering Research in Medicine

University of Bristol and United Bristol Healthcare Trust

Bristol General Hospital, Guinea Street, Bristol BS1 6SY, UK

{Mike.Halliwell,Peter.Wells}@bristol.ac.uk

## Abstract

*A new framework<sup>1</sup> for fusion of 2-D images based on their multiscale edges is described in this paper. The new method uses the multiscale edge representation of images proposed by Mallat and Hwang. The input images are fused using their multiscale edges only. Two different algorithms for fusing the point representations and the chain representations of the multiscale edges (wavelet transform modulus maxima) are given. The chain representation has been found to provide numerous new alternatives for image fusion, since edge graph fusion techniques can be employed to combine the images. The new framework encompasses different levels, i.e. pixel and feature levels, of image fusion in the wavelet domain.*

## 1. Image Fusion

The successful fusion of images acquired from different modalities or instruments is of great importance in many applications, such as medical imaging, microscopic imaging, remote sensing, computer vision and robotics. Image fusion can be defined as the process by which several images, or some of their features, are combined together to form a single image. Image fusion can be performed at different levels of the information representation. Four different levels can be distinguished according to [1], i.e. signal, pixel, feature and symbolic levels. To date, the results of image fusion in areas such as remote sensing and

medical imaging are primarily intended for presentation to a human observer for easier and enhanced interpretation. Therefore, the perception of the fused image is of paramount importance when evaluating different fusion schemes. When fusion is done at pixel level the input images are combined without any pre-processing. Pixel level fusion algorithms vary from very simple, e.g. image averaging, to very complex, e.g. principal component analysis (PCA), pyramid based image fusion and wavelet transform (WT) fusion. Several approaches to pixel level fusion can be distinguished, depending on whether the images are fused in the spatial domain or they are transformed into another domain, and their transforms are fused.

## 2. Wavelet Transform Fusion

Several wavelet based techniques for fusion of 2-D images have been described in the literature [3, 2, 4, 8]. In all wavelet based image fusion schemes the wavelet transforms  $\omega$  of the two registered input images  $I_1(x, y)$  and  $I_2(x, y)$  are computed and these transforms are combined using some fusion rule  $\phi$ . Then, the inverse wavelet transform  $\omega^{-1}$  is computed, and the fused image  $I(x, y)$  is reconstructed:

$$I(x, y) = \omega^{-1}(\phi(\omega(I_1(x, y)), \omega(I_2(x, y)))) \quad (1)$$

## 3. Pixel Based WT Fusion

The basic idea of all multiresolution fusion schemes is motivated by the fact that the human visual system is

<sup>1</sup>This work was funded by UK EPSRC Grant #GR/L53373.

primarily sensitive to local contrast changes, e.g. the edges or corners. In the case of pixel based wavelet transform fusion all respective wavelet coefficients from the input images are combined using the fusion rule  $\phi$ . Since wavelet coefficients having large absolute values contain the information about the salient features of the images such as edges and lines, a good fusion rule is to take the *maximum* of the (absolute values of the) corresponding wavelet coefficients. A more advanced *area based selection rule* is proposed in [4]. The maximum absolute value within a window is used as an activity measure of the central pixel of the window. A binary decision map of the same size as the WT is constructed to record the selection results based on a maximum selection rule. Another method called *contrast sensitivity fusion* is given in [8]. This method uses a weighted energy in the human perceptual domain, where the perceptual domain is based upon the frequency response, i.e. contrast sensitivity, of the human visual system. This wavelet transform image fusion scheme is an extension to the pyramid based scheme described by the same authors.

#### 4. Wavelets and Multiscale Edge Detection

The multiscale edge representation of images has evolved in two forms, based on the multiscale zero-crossings and multiscale gradient maxima, respectively. The latter approach, which is used in this study, was studied in details by Mallat and Zhong [7] and expanded to a tree-structured representation by Lu [5]. Here, we will briefly review this multiscale edge representation of two-dimensional images. The notation used is the same one as in [7]. Several important edge detection algorithms look for local maxima of the gradient of various smoothed versions of the image. This produces a hierarchy of edge features indexed by scale. Mallat and Zhong formalized the concept of multiscale edge representation with the wavelet transform associated with a particular class of non-orthogonal spline wavelets. The two-dimensional dyadic WT of an image  $I(x, y) \in L^2(\mathcal{R}^2)$  at scale  $2^j$  and in orientation  $k$  is defined as:

$$W_{2^j}^k f(x, y) = I * \psi_{2^j}^k(x, y), \quad k = 1, 2 \quad (2)$$

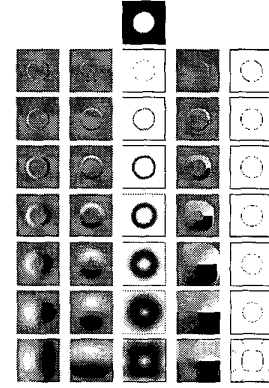
The two oriented wavelets  $\psi_{2^j}^k$  can be constructed by taking the partial derivatives

$$\psi^1(x, y) = \frac{\partial \theta(x, y)}{\partial x} \text{ and } \psi^2(x, y) = \frac{\partial \theta(x, y)}{\partial y} \quad (3)$$

where  $\theta(x, y)$  is a separable spline scaling function which plays the role of a smoothing filter. It can be shown that the two-dimensional WT defined by (2) gives the gradient of  $I(x, y)$  smoothed by  $\theta(x, y)$  at dyadic scales

$$\nabla_{2^j} I(x, y) \equiv (W_{2^j}^1 I(x, y), W_{2^j}^2 I(x, y)) =$$

$$= \frac{1}{2^{2j}} \nabla(\theta_{2^j} * I)(x, y) = \frac{1}{2^{2j}} \nabla I * \theta_{2^j}(x, y) \quad (4)$$



**Figure 1. A wavelet decomposition of a circle image.** The original image  $I(x, y)$  is on top. The five columns from left to right display: (a) the horizontal wavelet transform  $W_{2^j}^1 I(x, y)$ ; (b) the vertical wavelet transform  $W_{2^j}^2 I(x, y)$ ; (c) the wavelet transform modulus  $M_{2^j} I(x, y)$ ; (d) the wavelet transform angle  $A_{2^j} I(x, y)$  for a non-zero modulus; (e) the wavelet transform modulus maxima. The scale  $2^j$ , where  $j = 1, \dots, 7$ , increases from top to bottom.

If we want to locate the positions of rapid variation of an image  $I$ , we should consider the local maxima of the gradient magnitude at various scales which is given by

$$\begin{aligned} M_{2^j} I(x, y) &\equiv \|\nabla_{2^j} I(x, y)\| = \\ &= \sqrt{(W_{2^j}^1 I(x, y))^2 + (W_{2^j}^2 I(x, y))^2} \end{aligned} \quad (5)$$

A point  $(x, y)$  is a multiscale edge point at scale  $2^j$  if the magnitude of the gradient  $M_{2^j} I$  attains a local maximum there along the gradient direction  $A_{2^j} I$ , defined by

$$A_{2^j} I(x, y) \equiv \arctan \left[ \frac{W_{2^j}^2 I(x, y)}{W_{2^j}^1 I(x, y)} \right] \quad (6)$$

For each scale, we can collect the edge points together with the corresponding values of the gradient, i.e. the WT values, at that scale. The resulting local gradient maxima set at scale  $2^j$  is

$$P_{2^j}(I) = \{p_{2^j,i} = (x_i, y_i); \nabla_{2^j} I(x_i, y_i)\}, \quad (7)$$

where  $M_{2^j} I(x_i, y_i)$  has local maximum at  $p_{2^j,i} = (x_i, y_i)$  along the direction  $A_{2^j} I(x_i, y_i)$ . For a  $J$ -level two-dimensional dyadic WT, the set

$$\rho(I) = \{S_{2^j} I(x, y), [P_{2^j}(I)]_{1 \leq j < J}\} \quad (8)$$

is called a *multiscale edge representation of the image*  $I(x, y)$ . Here  $S_{2^J}I(x, y)$  is the low-pass approximation of  $I(x, y)$  at the coarsest scale  $2^J$ . The multiscale edge representation of a circle image is shown in Figure 1.

## 5. Edge Based WT Fusion

Mallat and Zhong proposed an algorithm in [7]<sup>2</sup> that reconstructs a very close and visually indistinguishable approximation of the input image from its multiscale edge representation. Thus, the multiscale edge representation of  $I$  is complete. The wavelets used by Mallat and Zhong are not orthogonal, but nevertheless,  $I$  may be recovered from its multiscale representation through the use of an associated family of synthesis wavelets (see [7] for details). This reconstruction algorithm is also stable for precisions of the order of 30 dB [7], which means that if the wavelet transform modulus maxima are slightly modified this will still lead to a very close approximation of the input image. By thresholding the wavelet transform modulus maxima based on their modulus values, it is possible to suppress the noise in the image, as well as some light textures. This kind of thresholding can be viewed as a nonlinear noise removal technique [7]. In [6] Mallat and Hwang have described a more complex method for suppression of white noise in images, whereby the maxima produced by the noise are removed based on their behaviour across scale.

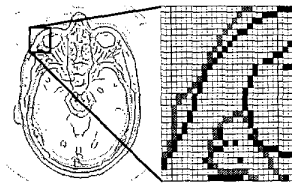


Figure 2. The point representations at scale  $2^3$  of the CT (grey edges) and MR (black edges) images of the head after thresholding.

### 5.1. Fusion of the Point Representations

As described in section 4, we can compute and store the resulting local gradient maxima sets  $P_{2^j}^1(I)$  and  $P_{2^j}^2(I)$  (refer to equation 7), which in this section will be also called the (edge) *point representation of the image*, for each one of the two input images and for each scale  $2^j$ .

<sup>2</sup>The authors would like to thank Stephane Mallat, Wen Hwang and Sifen Zhong, from New York University, for providing the *wave2* program.

Then, instead of combining all wavelet coefficients, we can *fuse the two wavelet transforms by combining their multiscale representations*, i.e. the local gradient maxima sets at each scale  $P_{2^j}^1(I)$  and  $P_{2^j}^2(I)$ , and the low-pass approximations  $S_{2^J}I_1(x, y)$  and  $S_{2^J}I_2(x, y)$  at the coarsest scale  $2^J$ . A possible fusion rule may be to take the union of the input point representations  $P_{2^j}(I) = P_{2^j}^1(I) \cup P_{2^j}^2(I)$ . In cases where an edge point  $(x_i, y_i)$  from  $I_1$  coincides spatially with an edge point  $(x_k, y_k)$  from  $I_2$ , the second edge point may overwrite the first, or vice versa. The low-pass approximations at the coarsest scale  $2^J$  may be fused pixel-wise by taking the average of the two subimages. Once the multiscale representation (see equation 8) of the fused image is constructed, the fused image itself can be reconstructed from it using the algorithm given in [7]. An example of image fusion of two medical images, where the point structures of the two images have been thresholded and then combined using the above-described fusion rule, is displayed in Figure 3. Only edge points with magnitude  $M_{2^j}I_1(x_i, y_i) > 2.0$  and  $M_{2^j}I_2(x_i, y_i) > 2.0$  are fused to create the point representation of the fused image.

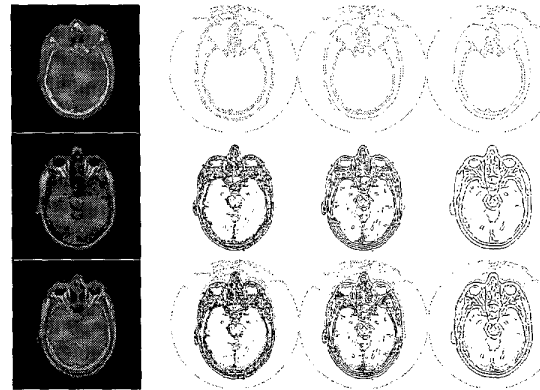
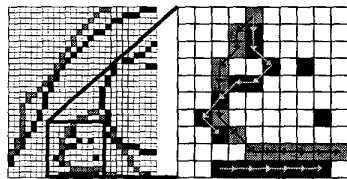


Figure 3. Fusion of CT (top row) and MR (middle row) images of a human head. The point representations of the two input images have been combined after thresholding (bottom row). The images are from the Visible Human Project dataset and are courtesy of the National Library of Medicine, USA.

### 5.2. Fusion of the Chain Representations

Sharp variations of 2-D images often belong to curves in the image plane [7]. Along these curves, the image intensity can be singular in one direction while varying smoothly in the perpendicular direction. It is known that

such curves are sometimes more meaningful than edge points by themselves, because they generally outline the boundaries of the image structures. In [7] Mallat and Zhong have created a high-level structure in the form of a collection of *chains of edge points*  $p_{2^j,i}$ , which here will be called the *chain representation of the image*. Two adjacent local maxima are chained together if their respective position is perpendicular to the direction indicated by  $A_{2^j}I(x,y)$ . Since the idea is to recover edge curves along which the image profile varies smoothly, maxima points are chained together only where the modulus  $M_{2^j}I(x,y)$  has close values. This chaining procedure defines a sparse, high-level image representation which consists of sets of maxima chains. The construction of the chain representation at one scale of the wavelet decomposition is illustrated in Figure 4.

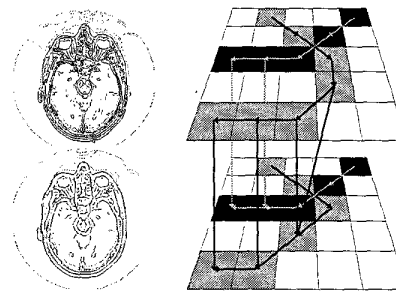


**Figure 4.** The point (left) and chain (right) representations at scale  $2^3$  of the CT (grey edges) and MR (black edges) images of the head (see also Figure 2). A close-up of the area around the left eye is displayed.

Such a high-level image representation turns out to be very useful in the image fusion process. Like in (5.1), the gradient maxima sets of the two input images have to be combined. In this case, however, much more sophisticated fusion rules can be employed. What we effectively have is *two oriented layered graphs*, where each layer corresponds to a certain scale  $2^j$  (see Figure 5). Hence, we can look for algorithms which will combine the two graphs and form one single graph. Distance measures calculating the proximity between nodes or edges of the two graphs can also be used in the decision process. Some nodes or edges can be merged and others can be removed in the output edge graph.

## 6. Conclusions

A new method for fusion of images, using their multiscale edges only, has been described in this study. The new method produces fusion results similar in quality to some other pixel based wavelet transform fusion methods proposed by other authors. The main difference, however,



**Figure 5.** The point representations after thresholding of a CT (grey edges) image and an MR (black edges) image at scale  $2^2$  (top left) and  $2^3$  (bottom left). The graphs (right) constructed by computing the chain representations of the images.

is that this is a feature based wavelet fusion method, which combines the high-level sparse representations of the input images in the form of multiscale edges (wavelet transform modulus maxima) or chains of such edge points, in order to fuse the images.

## References

- [1] M. A. Abidi and R. C. Gonzalez, editors. *Data Fusion in Robotics and Machine Intelligence*. Academic Press, 1992.
- [2] L. J. Chipman and T. M. Orr. Wavelets and image fusion. In *Proceedings IEEE International Conference on Image Processing, Washington D.C.*, pages 248–251. IEEE, Oct. 1995.
- [3] I. Koren, A. Laine, and F. Taylor. Image fusion using steerable dyadic wavelet transforms. In *Proceedings IEEE International Conference on Image Processing, Washington D.C.*, pages 232–235. IEEE, Oct. 1995.
- [4] H. Li, B. S. Manjunath, and S. K. Mitra. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57(3):235–245, May 1995.
- [5] J. Lu. *Signal recovery and noise reduction with wavelets*. PhD thesis, Dartmouth College, Hanover, New Hampshire, 1993.
- [6] S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, 38:617–643, Mar. 1992.
- [7] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):710–732, 1992.
- [8] T. A. Wilson, S. K. Rogers, and L. R. Myers. Perceptual based hyperspectral image fusion using multiresolution analysis. *Optical Engineering*, 34(11):3154–3164, 1995.